The Patent Office

INVESTOR IN PEOPLE

EP99/9284

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Andrew Conway

Dated 30 November 1999

Patents Form 1/77

ts Act 1977
16)

THE PATENT OFFICE
H
-1 DEC 1998
RECEIVED

The Patent Office

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1.	Your Reference	**MJS/MKR/PG3606**
2.	Patent application number *(The Patent office will fill in this part)*	01 DEC 1998 **9826359.3**
3.	Full name, address and postcode of the or of each applicant *(underline all surnames)*	**GLAXO GROUP LIMITED** **GLAXO WELLCOME HOUSE** **BERKELEY AVENUE** **GREENFORD** **MIDDLESEX** **UB6 ONN** **GB**
	Patents ADP number *(if you know it)*	07358700 3
	If the applicant is a corporate body, give the country/state of its corporation	**GB**
4	Title of the invention	**NOVEL RECEPTORS**
5	Name of your agent *(if you know one)*	**MICHAEL J STOTT (SEE CONTINUATION SHEET)**
	"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	**GLAXO WELLCOME PLC** **GLAXO WELLCOME HOUSE, BERKELEY AVENUE** **GREENFORD, MIDDLESEX** **UB6 ONN, GB**
	Patents ADP number *(if you know it)*	13375750 03

021DEC98 E408946-1 001030
P01/7700 0.00 - 9826359.3

6. If you are declaring priority from one or more earlier patent applications, give the country and date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number	Country	**Priority application number** *(if you know it)*	**Date of Filing** (day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	**Number of earlier application**	**Date of filing** (day / month / year)

8. Is a statement of inventorship and of right to grant a patent required in support of this request? *(Answer yes if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.

YES

See note (d))

Additional Agents
(See Page 1 No. 5)

NAME(S)

Alan HESKETH
Laurence David JENKINS
William Michael DADSON
Michael ATKINSON
Karen CRAWLEY
Peter I. DOLTON
Hugh B. DAWSON
Wendy Anne FILLER
Ruth Elizabeth HACKETT
Catriona MacLeod HAMMER
Audrey HAMMETT
Graham M.H. LANE
Stephanie Anne LEAROYD
Helen Kaye QUILLIN
Michael A REED
Marion REES
Michael John STOTT
Andrew J. TEUTEN
Rachel M. THORNLEY
Janis Florence VOLCKMAN

ADDRESS

Glaxo Wellcome plc
Glaxo Wellcome House
Berkeley Avenue
Greenford
Middlesex
UB6 ONN
Great Britain

Novel Receptors

Field of the Invention

The present invention relates to human vanilloid receptor (hVR) proteins and to related nucleotide sequences, expression vectors, cell lines, antibodies, screening methods, compounds, methods of production and methods of treatment, as well as other related aspects.

Background of the Invention

Capsaicin, the irritant in hot peppers and a member of the vanilloid family activates a sub-group of sensory neurons: the nociceptors. These neurons transmit nociceptive and thermoceptive pain information back to pain-processing centres in the central nervous system such as the spinal cord and the brain. They are also sites for the release of pro-inflammatory mediators in the periphery (1). Nociceptors show heterogeneity in their sensitivity to capsaicin. Excitation and prolonged exposure of these neurons to capsaicin is followed by a refractory state known as desensitisation (2) when they become insensitive to capsaicin and other noxious stimuli (3). The long-term response to insensitivity could be explained by death of the nociceptors or destruction of its peripheral terminals (4). Because of the desensitisation phenomenon, capsaicin has been used therapeutically for decades as an analgesic agent for the treatment of pain in a range of disorders (5).

It has been speculated that the endogenous target for capsaicin plays an important function in the detection of painful stimuli. It has been shown by electrophysiological and biochemical studies that capsaicin induces a flux of cations in dorsal root ganglion (DRG) neurons (6,7). Because other vanilloid derivatives show responses in a dose dependent manner (8,9) a receptor is the most likely candidate to explain the mechanism. Therefore, based on indirect

human vanilloid receptor sub-types could provide targets for the development of novel analgesic agents (agonists and antagonists) and agents which may interact with other disorders.

Accordingly, it is an object of the present invention to locate and characterise human vanilloid receptors. Other objects of the present invention will become apparent from the following detailed description thereof.

Summary of the Invention

According to one embodiment of the present invention there is provided an isolated human vanilloid receptor (hVR) protein or a variant thereof. Preferably the hVR protein is an hVR1, hVR2 or hVR3 protein or a variant thereof. In a particularly preferred aspect of the invention the hVR protein has an amino acid sequence as shown in Figure 2.

According to another aspect of the invention there is provided a nucleotide sequence encoding a human vanilloid receptor (hVR) protein or a variant thereof, or a nucleotide sequence which is complementary thereto. Preferably the nucleotide sequence encodes an hVR1, hVR2 or hVR3 protein or variant thereof or a nucleotide sequence which is complementary thereto. Particularly preferably the nucleotide sequence is as shown in Figure 1.

According to another aspect of the invention there is provided an expression vector comprising a nucleic acid sequence as referred to above which is capable of expressing an hVR protein.

According to another aspect of the invention there is provided a stable cell line comprising an expression vector as referred to above. Preferably the call line is a modified HEK293, C40 or HeLa cell line.

thereof, under conditions suitable for obtaining expression of the hVR protein or variant.

Brief Description of the Figures

Figure 1 shows the complete nucleotide sequence of the human VR1 including the 2517 open reading frame (base 51- 2568)

Figure 2 shows the nucleotide and encoded amino acid sequence of the human VR1 sequence.

Figure 3 shows the amino acid sequence of the human VR1 gene; the shading denotes predicted trans-membrane regions (boxed) and the ankyrin binding domains (unboxed). The predicted phosphorylation sites are underlined.

Figure 4 shows multiple alignment of the rat and human VR1 full-length amino acids sequences. The genes are described as: hVR1 for human VR1 and rVR1 for rat VR1.

Figure 5 shows a slot blot hybridisation of various tissues and clones with a probe derived from the amino acid coding region of the human VR1 gene.

Figure 6 shows a multiple alignment of the amino acids sequences of the various human vanilloid receptors genes identified with the rat VR1 sequence. They are described as: hVR1 for human VR1, rVR1 for rat VR1, 3primehVR2 for the 3' end of human VR2 5primehVR2 for the 5' end of human VR2 and 3endhVR3 for the 3' end of human VR3.

Figure 7 shows the nucleotide sequence of the 5' end of hVR2.

Figure 8 shows the nucleotide sequence of the 3' end of hVR2.

Routine methods, as further explained in the subsequent experimental section, can be employed to purify and/or synthesise the receptor proteins according to the invention. Such methods are well understood by persons skilled in the art, and include techniques such as those disclosed in Sambrook, J. et al (28), the disclosure of which is included herein in its entirety by way of reference.

By the term "variant" what is meant throughout the specification and claims is that other peptides or proteins which retain the same essential character of the vanilloid receptor proteins for which sequence information is provided, are also intended to be included within the scope of the invention. For example, other peptides or proteins with greater than about 80%, preferably at least 90% and particularly preferably at least 95% homology with the sequences provided are considered as variants of the receptor proteins. Such variants may include the deletion, modification or addition of single amino acids or groups of amino acids within the protein sequence, as long as the peptide maintains the biological functionality of a vanilloid receptor. This biological functionality can of course be assessed by conducting binding studies with the know vanilloid modulators: capsaicin, capsazepine (12) and resiniferatoxin (11). The term "variant" does, however, exclude the rat VR1 protein which has been previously identified (15).

Human VR1 is preferentially expressed in human dorsal root ganglia (DRG) and relative to hVR2 and hVR3 has the highest sequence homology with the rat VR1. Therefore, hVR1 is likely to be the human orthologue to rat VR1. hVR2 and hVR3 are less similar to rat VR1 and are expressed in a wider range of tissues. Nucleotide sequence analysis of hVR1 reveals a 2517bp open reading frame which encodes an 839 amino acid protein (see Figures 1, 2 and 3). This deduced protein sequence is 90 % identical to the rat VR1 (15) and shares many of its characteristics such as 6 transmembrane regions with an hydrophobic stretch between transmembrane 5 and 6 and an N-terminus which contains 3 ankyrin repeat domains.

which have been modified by insertion of vectors encoding for the receptor proteins according to the invention include the mammalian HEK293T, CHO, HeLa and COS cells. Preferably the cell line selected will be one which is not only stable, but also allows for mature glycosylation and cell surface expression of the inventive receptors.

It is also possible for the receptors of the invention to be transiently expressed in a cell line or on a membrane, such as for example in a baculovirus expression system. Such systems, which are adapted to express the receptors according to the invention, are also included within the scope of the present invention.

In particular, the functional hVR protein may include hVR receptor proteins selected from hVR1, hVR2 and hVR3 and thereof or even other hVR protein subtypes or splice variants which have not yet been identified.

According to another aspect, the present invention also relates to antibodies (either polyclonal or preferably monoclonal antibodies) which have been raised by standard techniques and are specific for the receptor proteins or variants thereof according to the invention. Such antibodies could for example, be useful in purification, isolation or screening involving immuno precipitation techniques and may be used as tools to further ellucidate hVR protein function, or indeed as therapeutic agents in their own right. Antibodies may also be raised against specific epitopes of the receptors according to the invention.

An important aspect of the present invention is the use of receptor proteins according to the invention in screening methods designed to identify compounds which act as receptor ligands and which may be useful to modulate receptor activity. In general terms, such screening methods will involve contacting the receptor protein concerned, preferably hVR1, hVR2 or hVR3, with a test compound and then detecting modulation in the receptor activity, or indeed

Pharmaceutical Sciences, Mack Publishing Company, Eastern Pennsylvania, 17th Ed, 1985, the disclosure of which is included herein in its entirety by way of reference.

The compounds may be administered via enteral or parenteral routes such as via oral, buccal, anal, pulmonary, intravenous, intraarterial, intramuscular, intraperitoneal, topical or other appropriate administration routes.

The present invention will be further explained, by way of example, in the appended experimental section.

Experimental

Identification of related human ESTs (Expressed Sequence Tags) (19) to the rat VR1 sequence by in silico analysis:

The full-length rat VR1 amino acid sequence (15) was used as a query sequence using the tBlastn (20) alignment program to identify related human genes in the dbEST (21) and Incyte (Palo Alto, Ca. USA) databases. Several human ESTs were identified and those with similarities greater than 50% selected for further analysis. One of these ESTs was T12251 previously shown to have 68% amino-acid identity and 84% similarity over a region of 70 amino acids (15). All human ESTs from both databases were clustered to identify overlapping identical ESTs belonging to the same transcript. The GCG package (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, Wisconsin) and a program developed in house termed ESTBlast (22) were used to build up these clusters. In total, forty-three ESTs derived from different tissue sources and both EST databases were clustered in ten groups named hVR2, hVR3, hVR1a, hVR1b, hVR1c, hVR1d, hVR1e, hVR1f, hVR1g and hVR1h. For each EST the tissue source was assigned according to the annotations in the dbEST and Incyte databases. The shortest cluster

hVR1c also while having high homology (71% identity and 82% similarity over 65 residues) was closely related to the C-terminus of the rat protein sequence.

Cluster hVR1e shared a high homology (67% identity and 78% similarity for amino acids) with the 5' of the rat VR1 sequence but did not seem to have a potential start codon. It contained two Incyte ESTs (Clones Ids: 3801964 and 3802764) derived from the same tissue, bladder, and from the same patient. These two ESTs were selected for further investigation since this cluster was the most 5', had high homology with rat VR1 and the bladder tissue could be contaminated with sensory neurons. Both cDNA clones were ordered but only 3801964 was received as 3802764 failed the recovering procedure established at Incyte laboratories. The cDNA insert was directionally cloned into the pINCY vector (Incyte).

The 3801964 clone was grown using standard procedures and DNA was isolated using Qiagen columns. SP6 (5' ATTTAGGTGACACTATAG) and T7 (5' TAATACGACTCACTATAGGG) primers flanking the cloning site of pINCY were used to sequence both ends. Plasmid DNA (0.6 pmol) was used with 10.0 pmol of each primer for the dye terminator reaction. The SP6 end corresponded to the *in silico* derived EST sequence (identical to 3801964). The T7 end did not have homologies with VR1 nor did it possess a long open reading frame or a polyadenylation motif (data not shown). The size of the insert was determined by enzyme digestion of the DNA with the endonucleases NotI and EcoRI and calculated to be approximately 3kb (data not shown).

Plasmid DNA (50ng) was used to amplify the insert by Polymerase Chain Reaction (PCR) with T7 and SP6 as primers. The PCR conditions included an initial hot-start at 94°C for 2 minutes, followed by 35 cycles at 94°C for 45 seconds, 50°C for 45 seconds and 72°C for 1 minute and terminated by 5 minutes at 72°C. The resulting

We formulated the hypothesis that both sequences (hVR1e and hVR3/1c) were part of a common transcript. If the human and rat VR1 were going to be similar, the 2 contigs should be separated by a gap of approximately 275bp. Primers were designed on both sides of the gap to amplify mRNA from DRG and brain tissues in order to clone the gap (Figure 6). A smear was obtained with the sense primer (5' TCTACTTCGGTGAACTGCCC) and antisense (5' ACGGCAGGGAGTCATTCTTC). For specificity 50ng of the PCR product were amplified with the nested sense (5' CTGCAGAACTCCTGGCAGA) and antisense (5' GTCACCACCGCTGTGGAAAA) primers. The 900bp nested amplicon was sequenced and shown to be identical to hVR1e at one end and hVR3/1c at the other end. The middle part of the PCR product was homologous to the rat VR1 sequence. This region corresponded to 91 amino acids. When the sequences of hVR1e, hVR3/hVR1c and the internal amplicon are combined the total length of the Open Reading Frame (ORF) is 823 amino acids followed by a 3' untranslated sequence of 1120 bp. The human amino acid sequence is 87% identical to the rat sequence over that part of the coding region. This sequence was termed hVR1 because of its high degree of identity with the rat VR1 sequence. It was a combination of the following clusters: HVR1c, hVR1e, hVR3.

Since no start codon was identified at the 5' end an additional strategy was designed to identify the full-length sequence. Two primers, sense (5' TCCTCTGGCTTCCAACCCGTT) and antisense (5' GAACTGGGCAGAAAGTGCCT) were designed to amplify a 150bp product from the first intron. A P1 Artificial Chromosome (PAC) genomic clone (25) was isolated by PCR screening of a PAC library (Genome Systems, St Louis, Missouri). PAC DNA was recovered by using standard plasmid isolation protocol (26). An anti-sense primer was designed (5' CTGGAGTTAGGGTCTCCATCC) to sequence the genomic clone towards the potential 5' end of the gene. An open reading frame with a starting codon was identified.

the hVR1 gene appears to be located on human chromosome 17 around marker SHGC-36073 (lod score=9.55).

In conclusion, we are confident that we have identified the full-length human orthologue of the rat VR1 gene because of the very high degree of homology with the rat transcript and a tissue distribution similar to that of rat VR1.

Identification and partial characterisation of additional human vanilloid receptors:

ESTs belonging to the remaining clusters were characterised by a combination of end sequencing, Bal-31 analysis and *in silico* cloning. The following Incyte clones were used during this process: -3324775 (hVR1a), -1725152 (hVR1b), -197476 (hVR1f), -3498406 (hVR1g), -1682513 (hVR1h), -1559496 (hVR2).

hVR2 was found to be a cluster of more than 40 ESTs derived from various tissues (and none from a DRG source). After *in silico* analysis and further sequencing of a particular Incyte clone, 1559496, two contiguous sequences were obtained: one located towards the 5' end and the other one more towards the 3' end. The former is 201 amino acid long while the latter has a size of 376 amino acids. A multiple comparison of these two contiguous sequences and the other vanilloid receptors is shown on figure 6. The 5' and 3' ends amino acid sequences of hVR2 have respectively 51% and 34% identity with the rat VR1 sequence.

A DNA probe from the coding region hVR2 was designed by enzyme restriction digest with Sty1 (New England Biolabs). A 438 bp fragment was obtained and used for radioactive hybridization on multi-Tissue northern blots (Clontech) according to the manufacturer's recommendations. As expected hVR2 is expressed in various tissues and a transcript of about 3.8 kb is detected by hybridization (figures 7 and 8).

shown by RT-PCR with the primer combination used to produce the probe that the gene is not expressed in DRG.

The 3' UTR sequence of hVR3 was used to design two primers to amplify a product of 360 bp: sense primer 5' ATGGCCACCAGCAGGGTTAC and antisense primer 5' TCTGCCAGGTTCCAGCTG. The G3 radiation hybrid panel from Stanford University (Research Genetics, Huntsville, Alabama) was screened by PCR. The positive lanes (eg: 11 lanes) and negative patterns were analysed using the public web server at Stanford University (http://www-sghc.stanford.edu). After analysis the hVR3 gene appears to be located on human chromosome 12 around markers D12S177E (lod score=15) and D12S1893 (lod score=14).

The same primer combination was used to isolate a genomic PAC clone (Genome Systems. St Louis, Missouri) to be used to identify the complete 5' end of the hVR3 transcript.

We have concluded that hVR3 is an additional human gene member of the vanilloid receptors family, but is not the orthologue of rVR1 because of its tissue distribution and homologies with rVR1.

hVR1g and hVR1b collapsed in a single contiguous sequence. Sequence analysis has shown that both cDNAs are likely to be chimeric. The 5' end has weak similarities with the rat VR1 gene but the 3' end is identical to a DNA binding protein.

Screening for compounds which exhibit hVR modulating activity

Mammalian cells, such as Hek293, CHO and HeLa cells over-expressing the VR receptor of choice are generated for use in the assay. 96 and 384 well plate, high throughput screens (HTS) are employed using fluorescence based calcium indicator molecules, including but not limited to dyes such as Fura-2, Fura-Red, Fluo 3 and Fluo 4 (Molecular Probes). Secondary screening involves

References

1. Szallasi, A. and Blumberg, P.M (1993) Mechanisms and therapeutic potential of vanilloids (capsaicin-like molecules). *Adv. Pharmacol.*, 24, 123-155.

2. Jansco, G. (1968) Desensitisation with capsaicin and related acylamides as a tool for studying the function of pain receptors. In: K. Lin, D. Armstrong and E.G. Pardo (Eds), Pharmacology of Pain, Pergamon Press, Oxford, 33-55.

3. Szolcsanyi, J. (1993) In *Capsaicin in the study of pain.* Ed. J. Wood, J. London: Academic Press, 1-26.

4. Szallasi, A. and Blumberg, P.M. (1996) Vanilloid receptors: new insights enhance potential as a therapeutic target. *Pain*, 68, 195-208.

5. Jansco, G., Kiraly, E. and Jansco-Gabor, A. (1977) Pharamacologically induced selective degeneration of chemosensitive primary sensory neurons. *Nature*, 270, 741-743.

6. Oh, U., Hwang, S.W. and Kim, D. (1996) Capsaicin activates a nonselective cation channel in cultured neonatal rat dorsal root ganglion neurons. *J. Neurosciences*, 16, 1659-1667.

7. Wood, J.N et al. (1988) Capsaicin-induced ion fluxes in dorsal root ganglion cells in culture. *J. Neurosciences*, 8, 3208-3220.

8. Szolcsanyi, J. and Jansco-Gabor, A. (1975) Sensory effects of capsaicin congeners I. Relationship between chemical structure and pain-producing potency of pungent agents. *Drug Res.*, 25, 1877-1881.

9. Szolcsanyi, J. and Jansco-Gabor, A. (1976) Sensory effects of capsaicin congeners II. Importance of chemical structure and pungency in desensitising activity of capsaicin-like compounds. Drug res., 26, 33-37.

10. James, I.F., Nikina, N. and Wood, J.N. (1993) The capsaicin receptor. In *Capsaicin in the Study of Pain.* Ed. Wood, J. London: Academic Press, 83-104.

11. Szallasi, A. and Blumberg, P.M. (1990) Specific binding of resiniferatoxin, an ultrapotent capsaicin analog, by dorsal root ganglion membranes. *Brain Research*, 524, 106-111.

of Expressed Sequence Tags (ESTs). *Computer Applications in Biosciences (CABIOS),* 13, 453-457.

23. Lau, P.P. and Gray H.B. (1979). *Nucleic Acids Research*, 6, 331.

24. Trower MK., Burt D., Purvis IJ., Dykes CW. & Christodoulou C. (1995). Fluorescent dye-primer cycle sequencing using non-purified PCR products as templates; development of a protocol amenable to high-throughput DNA sequencing. *Nucleic Acids Research* ,23, 2348-2349

25. Shepherd, N., Pfogner, B., Coulby, J., Ackerman, S., Vaidyanathan, G., Sauer, R., Balkenhol, T. and Sternberg, N. (1994). Preparation and screening of an arrayed human genomic library generated with the P1 cloning system. *Proc. Natl. Acad. Sci. USA.*, 91, 2629-2633.

26. Birboim, H.C. and Doly, J. (1979). *Nucl. Acids Res.*, 7, 1513-1523.

27. Burge, C. and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78-94.

28. Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular Cloning: a Laboratory Manual. 2nd Edition. CSH Laboratory Press. (1989)

10. The nucleotide sequence according to any one of claims 6 to 9 which is a cDNA sequence.

11. The nucleotide sequence according to claim 6 as shown in Figure 1 or Figure 9.

12. An expression vector comprising a nucleotide sequence according to any one of claims 6 to 11, which is capable of expressing an hVR protein.

13. The expression vector according to claim 12 capable of expressing a protein selected from hVR1, hVR2 or hVR3.

14. The expression vector according to claim 12 capable of expressing hVR1.

15. A stable cell line comprising an expression vector according to any one of claims 12 to 14.

16. The cell line according to claim 15 which is a modified HEK293, C40 or HeLa cell line.

17. An antibody specific for a protein as claimed in any one of claims 1 to 5.

18. A method for identification of a compound which exhibits hVR modulating activity comprising contacting an hVR protein according to any of claims 1 to 5 with a test compound and detecting modulating activity or inactivity.

19. A compound which modulates hVR activity, identifiable by a method according to claim 18, excluding the compounds capsaicin, resiniferatoxin and capsazepine.

Figure 1: hVR1 nucleotide sequence

```
   1  CGTGGTGGCT GCTGCAGGTT GCACACTGGG CCACAGAGGA TCCAGCAAGG
  51  ATGAAGAAAT GGAGCAGCAC AGACTTGGGG GCAGCTGCGG ACCCACTCCA
 101  AAAGGACACC TGCCCAGACC CCCTGGATGG AGACCCTAAC TCCAGGCCAC
 151  CTCCAGCCAA GCCCCAGCTC TCCACGGCCA AGAGCCGCAC CCGGCTCTTT
 201  GGGAAGGGTG ACTCGGAGGA GGCTTTCCCG GTGGATTGCC CTCACGAGGA
 251  AGGTGAGCTG GACTCCTGCC CGACCATCAC AGTCAGCCCT GTTATCACCA
 301  TCCAGAGGCC AGGAGACGGC CCCACCGGTG CCAGGCTGCT GTCCCAGGAC
 351  TCTGTCGCCG CCAGCACCGA GAAGACCCTC AGGCTCTATG ATCGCAGGAG
 401  TATCTTTGAA GCCGTTGCTC AGAATAACTG CCAGGATCTG GAGAGCCTGC
 451  TGCTCTTCCT GCAGAAGAGC AAGAAGCACC TCACAGACAA CGAGTTCAAA
 501  GACCCTGAGA CAGGGAAGAC CTGTCTGCTG AAAGCCATGC TCAACCTGCA
 551  CGACGGACAG AACACCACCA TCCCCCTGCT CCTGGAGATC GCGCGGCAAA
 601  CGGACAGCCT GAAGGAGCTT GTCAACGCCA GCTACACGGA CAGCTACTAC
 651  AAGGGCCAGA CAGCACTGCA CATCGCCATC GAGAGACGCA ACATGGCCCT
 701  GGTGACCCTC CTGGTGGAGA ACGGAGCAGA CGTCCAGGCT GCGGCCCATG
 751  GGGACTTCTT TAAGAAAACC AAAGGGCGGC CTGGATTCTA CTTCGGTGAA
 801  CTGCCCCTGT CCCTGGCCGC GTGCACCAAC CAGCTGGGCA TCGTGAAGTT
 851  CCTGCTGCAG AACTCCTGGC AGACGGCCGA CATCAGCGCC AGGGACTCGG
 901  TGGGCAACAC GGTGCTGCAC GCCCTGGTGG AGGTGGCCGA CAACACGGCC
 951  GACAACaCGA AGTTTGTgAC gAGCATGtaC AaTgAGATTC TGATCCTGGG
1001  GGCCAAACTG CAcCCGACGC TgAAGCTgGA GGAGCTCACC aACaAGAAGG
1051  GAATGACGCC GCTGGCTCTG GCAGCTGGGA CCgGGAAGAT CGGGGTCTTG
1101  GCCTATATTC TCCAGCGGGA GATCCAGGAG CCCGAGTGCA GGCACCTGTC
1151  CAGGAAGTTC ACCGAGTGgg cCTACGGGCC CGTGCACTCC TCGCTGTACG
1201  ACCTGTCCTG CATCGACACC TGCGAGAAGA ACTCGGTGCT GGAGGTGATC
1251  GCCTACAGCA GCAGCGAGAC CCCTAATCGC CACGACATGC TCTTGGTGGA
```

1301 GCCGCTGAAC CGACTCCTGC AGGACAAGTG GGACAGATTC GTCAAGCGCA

1351 TCTTCTACTT CAACTTCCTG GTCTACTGCC TGTACATGAT CATCTTCACC

1401 ATGGCTGCCT ACTACAGGCC CGTGGATGGC TTGCCTCCCT TTAAGATGGA

1451 AAAAACTGGA GACTATTTCC GAGTTACTGG AGAGATCCTG TCTGTGTTAG

1501 GAGGAGTCTA CTTCTTTTTC CGAGGGATTC AGTATTTCCT GCAGAGGCGG

1551 CCGTCGATGA AGACCCTGTT TGTGGACAGC TACAGTGAGA TGCTTTTCTT

1601 TCTGCAGTCA CTGTTCATGC TGGCCACCGT GGTGCTGTAC TTCAGCCACC

1651 TCAAGGAGTA TGTGGCTTCC ATGGTATTCT CCCTGGCCTT GGGCTGGACC

1701 AACATGCTCT ACTACACCCG CGGTTTCCAG CAGATGGGCA TCTATGCCGT

1751 CATGATAGAG AAGATGATCC TGAGAGACCT GTGCCGTTTC ATGTTTGTCT

1801 ACATCGTCTT CTTGTTCGGG TTTTCCACAG CGGTGGTGAC GCTGATTGAA

1851 GACGGGAAGA ATGACTCCCT GCCGTCTGAG TCCACGTCGC ACAGGTGGCG

1901 GGGGCCTGCC TGCAGGCCCC CCGATAGCTC CTACAACAGC CTGTACTCCA

1951 CCTGCCTGGA GCTGTTCAAG TTCACCATCG GCATGGGCGA CCTGGAGTTC

2001 ACTGAGAACT ATGACTTCAA GGCTGTCTTC ATCATCCTGC TGCTGGCCTA

2051 TGTAATTCTC ACCTACATCC TCCTGCTCAA CATGCTCATC GCCCTCATGG

2101 GTGAGACTGT CAACAAGATC GCACAGGAGA GCAAGAACAT CTGGAAGCTG

2151 CAGAGAGCCA TCACCATCCT GGACACGGAG AAGAGCTTCC TTAAGTGCAT

2201 GAGGAAGGCC TTCCGCTCAG GCAAGCTGCT GCAGGTGGGG TACACACCTG

2251 ATGGCAAGGA CGACTACCGG TGGTGCTTCA GGGTGGACGA GGTGAACTGG

2301 ACCACCTGGA ACACCAACGT GGGCATCATC AACGAAGACC CGGGCAACTG

2351 TGAkGGCGTC AAGCGCACCC TGAGCTTCTC CCTGCGGTCA AGCAGAGTTT

2401 CAGGCAGACA CTGGAAGAAC TTTGCCCTGG TCCCCCTTTT AAGAGAGGCA

2451 AGTGCTCGAG ATAGGCAGTC TGCTCAGCCC GAGGAAGTTT ATCTGCGACA

2501 GTTTTCAGGG TCTCTGAAGC CAGAGGACGC TGAGGTCTTC AAGAGTCCTG

2551 CCGCTTCCGG GGAGAAGTGA GGACGTCACG CAGACAGCAC TGTCAACACT

2601 GGGCCTTAGG AGACCCCGTT GCCACGGGGX XCTGCTGAGG GAACACCAGT

2651 GCTCTGTCAG CAGCCTGGCC TGGTCTGTGC CTGCCCAGCA TGTTCCCAAA

2701 TCTGTGCTGG ACAAGCTGTG GGAAGCGTTC TTGGAAGCAT GGGGAGTGAT

2751 GTACATCCAA CCGTCACTGT CCCCAAGTGA ATCTCCTAAC AGACTTTCAG

2801 GTTTTTACTC ACTTTACTAA ACAGTTTGGA TGGTCAGTCT CTACTGGGAC

2851 ATGTTAGGCC CTTGTTTTCT TTGATTTTAT TCTTTTCTGT GAGACAGAGT

2901 TCACTCTTGT TGCCCAGGCT GGAGTGCAGT GGTGTGATCT TGGCTCACTG

2951 CAACCTCTGC TCCCGGGTTC AAGCGATTCT TCTGCTTCAG TCTCCCAAGT

3001 AGCTTGGATT ACAGGTGAGC ACTACCACGC CCGGCTAATT TTTGTATTTT

3051 TAATAGAGAC GGGGTTTCAC CATGTTGGCC AGGCTGGTCT CGAACTCTTG

3101 ACCTCAGGTG ATCTGCCCGC CTTGGCCTCC CAAAGTGCTG GGATTACAGG

3151 TGTGAGCCGC TGCGCTCGGC CTTCTTTGAT TTTATATTAT TAGGAGCAAA

3201 AGTAAATGAA GCCCAGGAAA ACACCTTTGG GAACAAACTC TTCCTTTGAT

3251 GGAAAATGCA GAGGCCCTTC CTCTCTGTGC CGTGCTTGCT CCTCTTACCT

3301 GCCCGGGTGG TTTGGGGGTG TTGGTGTTTC CTCCCTGGAG AAGATGGGGG

3351 AGGCTGTCCC ACTCCCAGCT CTGGCAGAAT CAAGCTGTTG CAGCAGTGCC

3401 TTCTTCATCC TTCCTTACGA TCAATCACAG TCTCCAGAAG ATCAGCTCAA

3451 TTGCTGTGCA GGTTAAAACT ACAGAACCAC ATCCCAAAGG TACCTGGTAA

3501 GAATGTTTGA AAGATCTTCC ATTTCTAGGA ACCCCAGTCC TGCTTCTCCG

3551 CAATGGCACA TGCTTCCACT CCATCCATAC TGGCATCCTC AAATAAACAG

3601 ATATGTATAC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A

Figure 2: Nucleotide and amino acid sequence of hVR1

```
-50 cgtggtggctgctgcaggttgcacactgggccacagaggatccagcaaggATGAAGAAAT 10
                                                      M  K  K  W  4

 11 GGAGCAGCACAGACTTGGGGGCAGCTGCGGACCCACTCCAAAAGGACACCTGCCCAGACC 70
  5   S  S  T  D  L  G  A  A  A  D  P  L  Q  K  D  T  C  P  D  P 24

 71 CCCTGGATGGAGACCCTAACTCCAGGCCACCTCCAGCCAAGCCCCAGCTCTCCACGGCCA 130
 25   L  D  G  D  P  N  S  R  P  P  P  A  K  P  Q  L  S  T  A  K 44

131 AGAGCCGCACCCGGCTCTTTGGGAAGGGTGACTCGGAGGAGGCTTTCCCGGTGGATTGCC 190
 45   S  R  T  R  L  F  G  K  G  D  S  E  E  A  F  P  V  D  C  P 64

191 CTCACGAGGAAGGTGAGCTGGACTCCTGCCCGACCATCACAGTCAGCCCTGTTATCACCA 250
 65   H  E  E  G  E  L  D  S  C  P  T  I  T  V  S  P  V  I  T  I 84

251 TCCAGAGGCCAGGAGACGGCCCCACCGGTGCCAGGCTGCTGTCCCAGGACTCTGTCGCCG 310
 85   Q  R  P  G  D  G  P  T  G  A  R  L  L  S  Q  D  S  V  A  A 104

311 CCAGCACCGAGAAGACCCTCAGGCTCTATGATCGCAGGAGTATCTTTGAAGCCGTTGCTC 370
105   S  T  E  K  T  L  R  L  Y  D  R  R  S  I  F  E  A  V  A  Q 124

371 AGAATAACTGCCAGGATCTGGAGAGCCTGCTGCTCTTCCTGCAGAAGAGCAAGAAGCACC 430
125   N  N  C  Q  D  L  E  S  L  L  L  F  L  Q  K  S  K  K  H  L 144

431 TCACAGACAACGAGTTCAAAGACCCTGAGACAGGGAAGACCTGTCTGCTGAAAGCCATGC 490
145   T  D  N  E  F  K  D  P  E  T  G  K  T  C  L  L  K  A  M  L 164

491 TCAACCTGGACGACGGACAGAACACCACCATCCCCCTGCTCCTGGAGATCGCGCGGCAAA 550
165   N  L  H  D  G  Q  N  T  T  I  P  L  L  L  E  I  A  R  Q  T 184

551 CGGACAGCCTGAAGGAGGTTGTCAACGCCAGCTACACGGACAGCTACTACAAGGGCCAGA 610
185   D  S  L  K  E  L  V  N  A  S  Y  T  D  S  Y  Y  K  G  Q  T 204

611 CAGCACTGCACATGCCATCGAGAGACGCAACATGGCCCTGGTGACCCTCCTGGTGGAGA 670
205   A  L  H  I  A  I  E  R  R  N  M  A  L  V  T  L  L  V  E  N 224

671 ACGGAGCAGACGTCCAGGCTGCGGCCCATGGGGACTTCTTTAAGAAAACCAAAGGGCGGC 730
225   G  A  D  V  Q  A  A  A  H  G  D  F  F  K  K  T  K  G  R  P 244

731 CTGGATTCTACTTCGGTGAACTGCCCCTGTCCCTGGCCGCGTGCACCAACCAGCTGGGCA 790
245   G  F  Y  F  G  E  L  P  L  S  L  A  A  C  T  N  Q  L  G  I 264

791 TCGTGAAgTTCCTGCTGCAGAACTCCTGGCAGACGGCCGACATCAGCGCCAGGGACTCGG 850
265   V  K  F  L  L  Q  N  S  W  Q  T  A  D  I  S  A  R  D  S  V 284

851 TGGGCAACACGGTGCTGCACGCCCTGGTGGAGGTGGCCGACAACaCGGCCGACAACaCGA 910
285   G  N  T  V  L  H  A  L  V  E  V  A  D  N  T  A  D  N  T  K 304

911 AGTTTGTgACgAGCATGtaCAaTgAGATTCTGATCCTGGGGGCCAAACTGCAcCCGACGC 970
305   F  V  T  S  M  Y  N  E  I  L  I  L  G  A  K  L  H  P  T  L 324

971 TgAAGCTgGAGGAGCTCACCaACaAGAAGGGAATGACGCCGCTGGCTCTGGCAGCTGGGA 1030
325   K  L  E  E  L  T  N  K  K  G  M  T  P  L  A  L  A  A  G  T 344

1031 CCgGGAAGATCGGGGTCTTGGCCTATATTCTCCAGCGGGAGATCCAGGAGCCCGAGTGCA 1090
 345   G  K  I  G  V  L  A  Y  I  L  Q  R  E  I  Q  E  P  E  C  R 364

1091 GGCACCTGTCCAGGAAGTTCACCGAGTGggcCTACGGGCCCGTGCACTCCTCGCTGTACG 1150
 365   H  L  S  R  K  F  T  E  W  A  Y  G  P  V  H  S  S  L  Y  D 384

1151 ACCTGTCCTGCATCGACACCTGCGAGAAGAACTCGGTGCTGGAGGTGATCGCCTACAGCA 1210
 385   L  S  C  I  D  T  C  E  K  N  S  V  L  E  V  I  A  Y  S  S 404

1211 GCAGCGAGACCCCTAATCGCCACGACATGCTCTTGGTGGAGCCGCTGAACCGACTCCTGC 1270
 405   S  E  T  P  N  R  H  D  M  L  L  V  E  P  L  N  R  L  L  Q 424

1271 AGGACAAGTGGGACAGATTCGTCAAGCGCATCTTCTACTTCAACTTCCTGGTCTACTGCC 1330
```

```
 425   D  K  W  D  R  F  V  K  R  I  F  Y  F  N  F  L  V  Y  C  L 444

1331 TGTACATGATCATCTTCACCATGGCTGCCTACTACAGGCCCGTGGATGGCTTGCCTCCCT 1390
 445   Y  M  I  I  F  T  M  A  A  Y  Y  R  P  V  D  G  L  P  P  F 464

1391 TTAAGATGGAAAAAACTGGAGACTATTTCCGAGTTACTGGAGAGATCCTGTCTGTGTTAG 1450
 465   K  M  E  K  T  G  D  Y  F  R  V  T  G  E  I  L  S  V  L  G 484

1451 GAGGAGTCTACTTCTTTTTCCGAGGGATTCAGTATTTCCTGCAGAGGCGGCCGTCGATGA 1510
 485   G  V  Y  F  F  F  R  G  I  Q  Y  F  L  Q  R  R  P  S  M  K 504

1511 AGACCCTGTTTGTGGACAGCTACAGTGAGATGCTTTTCTTTCTGCAGTCACTGTTCATGC 1570
 505   T  L  F  V  D  S  Y  S  E  M  L  F  F  L  Q  S  L  F  M  L 524

1571 TGGCCACCGTGGTGCTGTACTTCAGCCACCTCAAGGAGTATGTGGCTTCCATGGTATTCT 1630
 525   A  T  V  V  L  Y  F  S  H  L  K  E  Y  V  A  S  M  V  F  S 544

1631 CCCTGGCCTTGGGCTGGACCAACATGCTCTACTACACCCGCGGTTTCCAGCAGATGGGCA 1690
 545   L  A  L  G  W  T  N  M  L  Y  Y  T  R  G  F  Q  Q  M  G  I 564

1691 TCTATGCCGTCATGATAGAGAAGATGATCCTGAGAGACCTGTGCCGTTTCATGTTTGTCT 1750
 565   Y  A  V  M  I  E  K  M  I  L  R  D  L  C  R  F  M  F  V  Y 584

1751 ACATCGTCTTCTTGTTCGGGTTTTCCACAGCGGTGGTGACGCTGATTGAAGACGGGAAGA 1810
 585   I  V  F  L  F  G  F  S  T  A  V  V  T  L  I  E  D  G  K  N 604

1811 ATGACTCCCTGCCGTCTGAGTCCACGTCGCACAGGTGGCGGGGGCCTGCCTGCAGGCCCC 1870
 605   D  S  L  P  S  E  S  T  S  H  R  W  R  G  P  A  C  R  P  P 624

1871 CCGATAGCTCCTACAACAGCCTGTACTCCACCTGCCTGGAGCTGTTCAAGTTCACCATCG 1930
 625   D  S  S  Y  N  S  L  Y  S  T  C  L  E  L  F  K  F  T  I  G 644

1931 GCATGGGCGACCTGGAGTTCACTGAGAACTATGACTTCAAGGCTGTCTTCATCATCCTGC 1990
 645   M  G  D  L  E  F  T  E  N  Y  D  F  K  A  V  F  I  I  L  L 664

1991 TGCTGGCCTATGTAATTCTCACCTACATCCTCCTGCTCAACATGCTCATCGCCCTCATGG 2050
 665   L  A  Y  Y  V  I  L  T  Y  I  L  L  L  N  M  L  I  A  L  M  G 684

2051 GTGAGACTGTCAACAAGATCGCACAGGAGAGCAAGAACATCTGGAAGCTGCAGAGAGCCA 2110
 685   E  T  V  N  K  I  A  Q  E  S  K  N  I  W  K  L  Q  R  A  I 704

2111 TCACCATCCTGGACACGGAGAAGAGCTTCCTTAAGTGCATGAGGAAGGCCTTCCGCTCAG 2170
 705   T  I  L  D  T  E  K  S  F  L  K  C  M  R  K  A  F  R  S  G 724

2171 GCAAGCTGCTGCAGGTGGGGTACACACCTGATGGCAAGGACGACTACCGGTGGTGCTTCA 2230
 725   K  L  L  Q  V  G  Y  T  P  D  G  K  D  D  Y  R  W  C  F  R 744

2231 GGGTGGACGAGGTGAACTGGACCACCTGGAACACCAACGTGGGCATCATCAACGAAGACC 2290
 745   V  D  E  V  N  W  T  T  W  N  T  N  V  G  I  I  N  E  D  P 764

2291 CGGGCAACTGTGAkGGCGTCAAGCGCACCCTGAGCTTCTCCCTGCGGTCAAGCAGAGTTT 2350
 765   G  N  C  ?  G  V  K  R  T  L  S  F  S  L  R  S  S  R  V  S 784

2351 CAGGCAGACACTGGAAGAACTTTGCCCTGGTCCCCCTTTTAAGAGAGGCAAGTGCTCGAG 2410
 785   G  R  H  W  K  N  F  A  L  V  P  L  L  R  E  A  S  A  R  D 804

2411 ATAGGCAGTCTGCTCAGCCCGAGGAAGTTTATCTGCGACAGTTTTCAGGGTCTCTGAAGC 2470
 805   R  Q  S  A  Q  P  E  E  V  Y  L  R  Q  F  S  G  S  L  K  P 824

2471 CAGAGGACGCTGAGGTCTTCAAGAGTCCTGCCGCTTCCGGGGAGAAGTGAggacgtcacg 2530
 825   E  D  A  E  V  F  K  S  P  A  A  S  G  E  K  *              840

2531 cagacagcactgtcaacactgggccttaggagaccccgttgccacggggxxctgctgagg 2590

2591 gaacaccagtgctctgtcagcagcctggcctggtctgtgcctgcccagcatgttcccaaa 2650

2652 tctgtgctggacaagctgtgggaagcgttcttggaagcatggggagtgatgtacatccaa 2710

2711 ccgtcactgtccccaagtgaatctcctaacagactttcaggtttttactcactttactaa 2770

2771 acagtttggatggtcagtctctactgggacatgttaggcccttgttttctttgattttat 2830
```

2831 tcttttctgtgagacagagttcactcttgttgcccaggctggagtgcagtggtgtgatct 2890

2891 tggctcactgcaacctctgctcccgggttcaagcgattcttctgcttcagtctcccaagt 2950

2951 agcttggattacaggtgagcactaccacgcccggctaatttttgtatttttaatagagac 3010

3011 ggggtttcaccatgttggccaggctggtctcgaactcttgacctcaggtgatctgcccgc 3070

3071 cttggcctcccaaagtgctgggattacaggtgtgagccgctgcgctcggccttctttgat 3130

3131 tttatattattaggagcaaaagtaaatgaagcccaggaaaacacctttgggaacaaactc 3190

3191 ttcctttgatggaaaatgcagaggcccttcctctctgtgccgtgcttgctcctcttacct 3250

3251 gcccgggtggtttgggggtgttggtgtttcctccctggagaagatgggggaggctgtccc 3310

3311 actcccagctctggcagaatcaagctgttgcagcagtgccttcttcatccttccttacga 3370

3371 tcaatcacagtctccagaagatcagctcaattgctgtgcaggttaaaactacagaaccac 3430

3431 atcccaaaggtacctggtaagaatgtttgaaagatcttccatttctaggaaccccagtcc 3490

3491 tgcttctccgcaatggcacatgcttccactccatccatactggcatcctcaaataaacag 3550

3551 atatgtatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3591

Figure 3: Amino acid sequence of hVR1

```
  1  MKKWSSTDLG AAADPLQKDT CPDPLDGDPN SRPPPAKPQL STAKSRTRLF
 51  GKGDSEEAFP VDCPHEEGEL DSCPTITVSP VITIQRPGDG PTGARLLSQD
101  SVAASTEKTL RLYDRRSIFE AVAQNNCQDL ESLLLFLQKS KKHLTDNEFK
151  DPETGKTCLL KAMLNLHDGQ NTTIPLLLEI ARQTDSLKEL VNASYTDSYY
201  [Ankyrin binding domains]GDFFKKT KGRPGFY[Ankyrin binding domains]
251  [Ankyrin binding domains] RDSVGNTVLH ALVEVADNTA
301  DNTKFVTSMY NEILILGAKL HPTLKLEELT NK[Ankyrin binding domains]
351  [Ankyrin binding domains]LSRKF TEWAYGPVHS SLYDLSCIDT CEKNSVLEVI
401  AYSSSETPNR HDMLLVEPLN RLLQDKWDRF VKR[Transmembrane domains]
451  [Transmembrane domains]RPVDG LPPFKMEKTG DYFRVTGEI[Transmembrane domains]FLQRR
501  PSMKTLFVD[Transmembrane domains]HLKEYVAS MV[Transmembrane domains]
551  [Transmembrane domains]E KMILRD[Transmembrane domains]TLIE
601  DGKNDSLPSE STSHRWRGPA CRPPDSSYNS LYSTCLELFK FTIGMGDLEF
651  TENYD[Transmembrane domains]ETVNKI AQESKNIWKL
701  QRAITILDTE KSFLKCMRKA FRSGKLLQVG YTPDGKDDYR WCFRVDEVNW
751  TTWNTNVGII NEDPGNCXGV KRTLSFSLRS SRVSGRHWKN FALVPLLREA
801  SARDRQSAQP EEVYLRQFSG SLKPEDAEVF KSPAASGEK*
```

Key

T/S predicted phosphorylation sites

[Transmembrane domains] Transmembrane domains

[Ankyrin binding domains] Ankyrin binding domains

Figure 4: Comparison of the amino acid sequences of the rat (rVR1) and human (hVR1) vanilloid proteins.

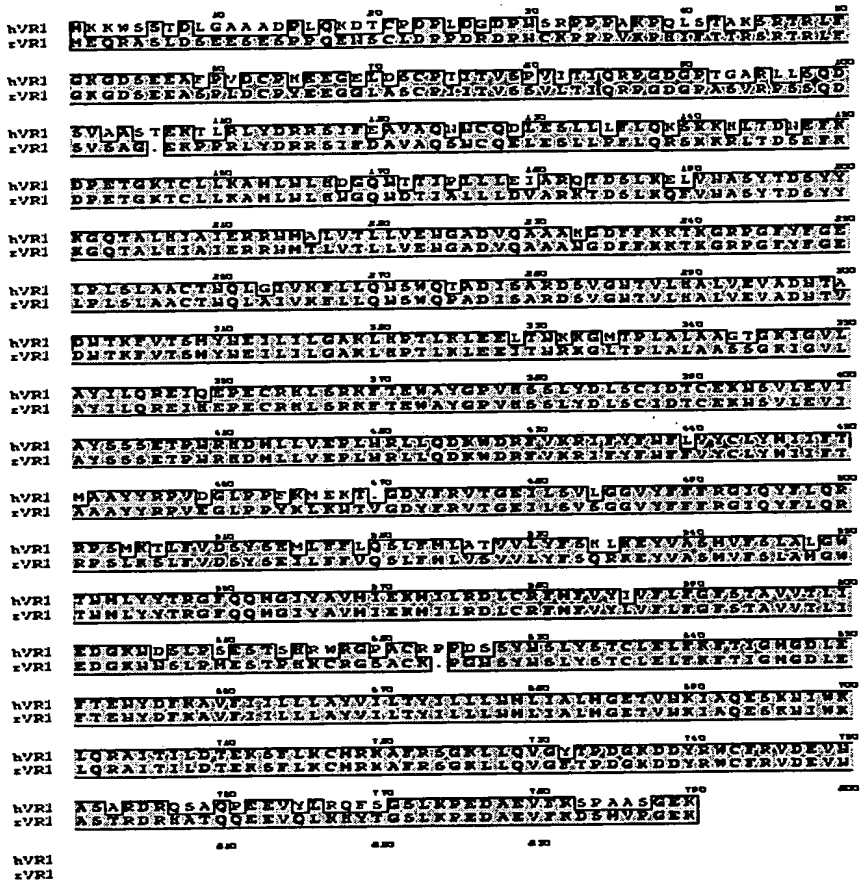
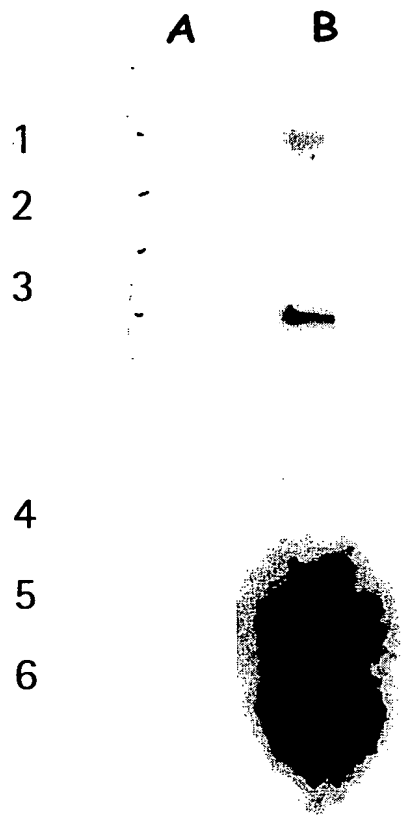


Figure 5: Slot Blot hybridisation with hVR1 probe

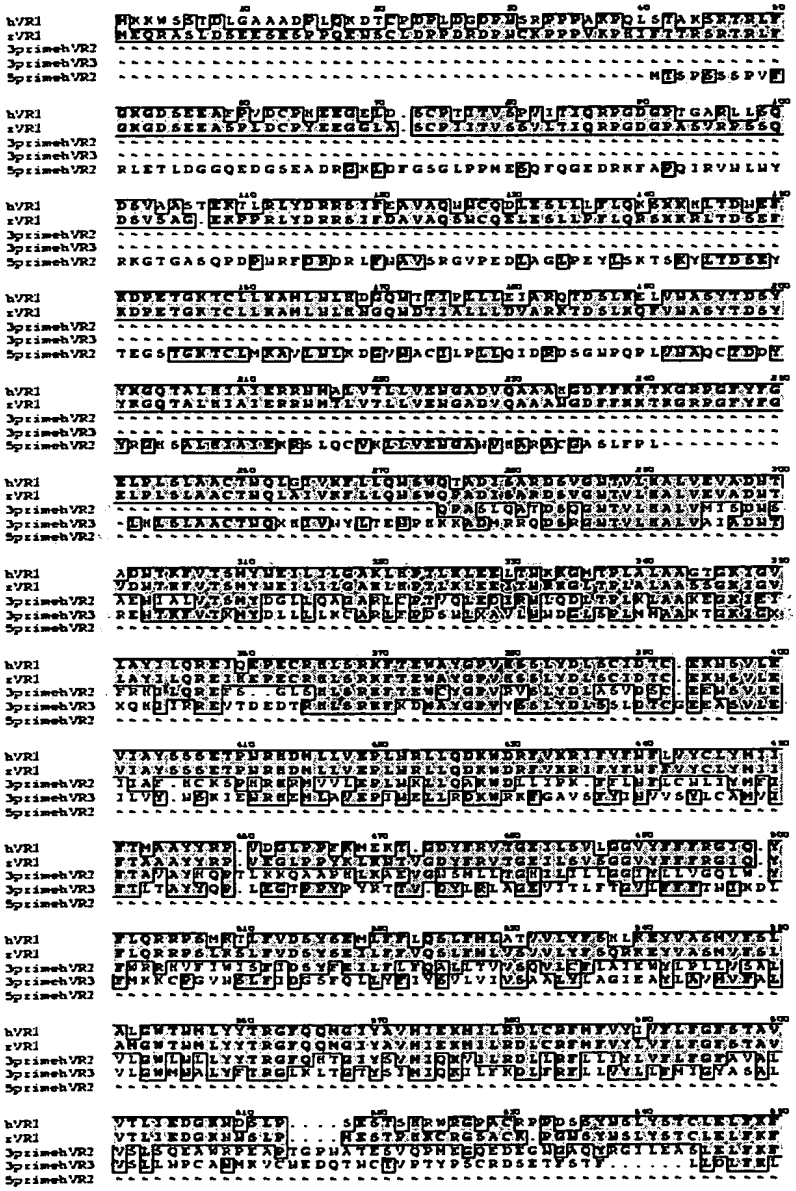


Well

1A hDRG
1B hDRG
2A rDRG
3A rat Spinal cord
3B hSpinal Cord

4B Water
5B Incyte clone 1427927
6B 900bp Amplicon from Brain cDNA

Figure 6: Multiple comparison of the amino acid sequ nces of rat VR1 and the human vanilloid receptors: hVR1, the 5' and 3' ends of hVR2 and the 3' end of hVR3.



igure 7

```
   1  ACCGGCGAAC GTGCNAGAAG GAAGGAAGAA AGCAAAGAAC GGCGTAGGGC
  51  GCTGGCAAGT GTAGCGGTCA CGCTGCGCGT AACCACCACA CCCGCCGCGC
 101  TTAATGCGCC GCTACAGGGC GCGTCCCATT CGCCATTCAG GCTGCGCAAC
 151  TGTTGGGAAG GGCGATCGGT GCGGGCCTCT TCGCTATTAC GCCAGCTGGC
 201  GAAAGGGGGA TGTGCTGCAA GGCGATTAAG TTGGGTAACG CCAGGGTTTT
 251  CCCAGTCACG ACGTTGTAAA ACGACGGCCA GTGAATTGAA TTTAGGTGAC
 301  ACTATAGAAG AGCTATGACG TCGCATGCAC GCGTACGTAA GCTCGGAATT
 351  CGGCTCGAGC CCACACCAGC CCGCCAGCCT GCAGGCCACT GACTCCCAGG
 401  GCAACACAGT CCTGCATGCC CTAGTGATGA TCTCGGACAA CTCAGCTGAG
 451  AACATTGCAC TGGTGACCAG CATGTATGAT GGGCTCCTCC AAGCTGGGGC
 501  CCGCCTCTGC CCTACCGTGC AGCTTGAGGA CATCCGCAAC CTGCAGGATC
 551  TCACGCCTCT GAAGCTGGCC GCCAAGGAGG GCAAGATCGA GATTTTCAGG
 601  CACATCCTGC AGCGGGAGTT TTCAGGACTG AGCCACCTTT CCCGAAAGTT
 651  CACCGAGTGG TGCTATGGGC CTGTCCGGGT GTCGCTGTAT GACCTGGCTT
 701  CTGTGGACAG CTGTGAGGAG AACTCAGTGC TGGAGATCAT TGCCTTTCAT
 751  TGCAAGAGCC CGCACCGACA CCGAATGGTC GTTTTGGAGC CCCTGAACAA
 801  ACTGCTGCAG GCGAAATGGG ATCTGCTCAT CCCCAAGTTC TTCTTAAACT
 851  TCCTGTGTAA TCTGATCTAC ATGTTCATCT TCACCGCTGT TGCCTACCAT
 901  CAGCCTACCC TGAAGAAGCA GGCCGCCCCT CACCTGAAAG CGGAGGTTGG
 951  AAACTCCATG CTGCTGACGG GCCACATCCT TATCCTGCTA GGGGGGATCT
1001  ACCTCCTCGT GGGCCAGCTG TGGTACTTCT GGCGGCGCCA CGTGTTCATC
1051  TGGATCTCGT TCATAGACAG CTACTTTGAA ATCCTCTTCC TGTTCCAGGC
1101  CCTGCTCACA GTGGTGTCCC AGGTGCTGTG TTTCCTGGCC ATCGAGTGGT
1151  ACCTGCCCCT GCTTGTGTCT GCGCTGGTGC TGGGCTGGCT GAACCTGCTT
1201  TACTATACAC GTGGCTTCCA GCACACAGGC ATCTACAGTG TCATGATCCA
1251  GAAGGTCATC CTGCGGGACC TGCTGCGCTT CCTTCTGATC TACTTAGTCT
1301  TCCTTTTCGG CTTCGCTGTA GCCCTGGTGA GCCTGAGCCA GGAGGCTTGG
```

1351 CGCCCCGAAG CTCCTACAGG CCCCAATGCC ACAGAGTCAG TGCAGCCCAT

1401 GGAGGGACAG GAGGACGAGG GCAACGGGGC CCAGTACAGG GGTATCCTGG

1451 AAGCCTCCTT GGAGCTCTTC AAATTCACCA TCGGCATGGG CGAGCTGGCC

1501 TTCCAGGAGC AGCTGCACTT CCGCGGCATG GTGCTGCTGC TGCTGCTGGC

1551 CTACGTGCTG CTCACCTACA TCCTGCTGCT CAACATGCTC ATCGCCCTCA

1601 TGAGCGAGAC CGTCAACAGT GTCGCCACTG ACAGCTGGAG CATCTGGAAG

1651 CTGCAGAAAG CCATCTCTGT CCTGGAGATG GAGAATGGCT ATTGGTGGTG

1701 CAGGAAGAAG CAGCGGGCAG GTGTGATGCT GACCGTTGGC ACTAAGCCAG

1751 ATGGCAGCCC CGATGAGCGC TGGTGCTTCA GGGTGGAGGA GGTGAACTGG

1801 GCTTCATGGG AGCAGACGCT GCCTACGCTG TGTGAGGACC CGTCAGGGGC

1851 AGGTGTCCCT CGTGAGTAGC CTGGAACTCT CGAGAACCCT GTCCTGGCTT

1901 CCCCTCCCAA GGAGGATGAG GATGGTGCCT CTGAGGAAAA CTATGTGCCC

1951 GTCCAGCTCC TCAGTCCAAC TGATGGCCCA GATGCAGCAG GAGGCCAGAG

2001 GACAGAGCAG AGGATCTTTC CAACCACATC TGCTGGCTCT GGGGTCCCAG

2051 TGAATTCTGG TGGCAAATAT ATATTTTCAC TAACTAAAAA AAAAAAAAAA

igure 8

```
   1  GGGAAGAATC CCCATCnATG GCAGCTTCCA TGGGTGGCAA GTCCCCAGCA
  51  TCCAAGGGCT GCCTCTGAGn GTCACCCACC CCCACCTGAG ACCTTAGTGG
 101  CTAGAATnnG GAnGGnTGGn GGTGGAnCCT nAnTCGCAGC AGGGTGTGTC
 151  CAGATGGTCA GTCTCTGGTG GCTAGCCTGT CCTGACAGGG GAGAGTTAAG
 201  CTCCCGtTCT CCACCGTGCC GGCTGGCaGG TGGGCTGAGG GTGACCGAGA
 251  GACCAGAACC TGCTTGCTGG AGCTTAGTGC TCAGAGCTGG GGAGGGAGGT
 301  TCCGCCGCTC CTCTGCTGTC AGCGCCGGCA GCCCCTCCCG GCTTCACTTC
 351  CTCCCGCAGC CCCTGCTACT GAGAAGCTCC GGGATCCCAG CAGCCGCCAC
 401  GCCCTGGCCT CAGCCTGCGG GGCTCCAGTC AGGCCAACAC CGACGCGCAn
 451  CTGGnGAGGA AGACAGGACC CTTGACATCT CCATCTGCAC AGAGGTCCTG
 501  GCTGGACCGA GCAGCCTCCT CCTCCTAGGA TGACCTCACC CTCCAGCTCT
 551  CCAGTTTTCA GGTTGGAGAC ATTAGATGGA GGCCAAGAAG ATGGCTCTGA
 601  GGCGGACAGA GGAAAGCTGG ATTTTGGGAG CGGGCTGCCT CCCATGGAGT
 651  CACAGTTCCA GGGCGAGGAC CGGAAATTCG CCCCTCAGAT AAGAGTCAAC
 701  CTCAACTACC GAAAGGGAAC AGGTGCCAGT CAGCCGGATC CAAACCGATT
 751  TGACCGAGAT CGGCTCTTCA ATGCGGTCTC CCGGGGTGTC CCCGAGGATC
 801  TGGCTGGACT TCCAGAGTAC CTGAGCAAGA CCAGCAAGTA CCTCACCGAC
 851  TCGGAATACA CAGAGGGCTC CACAGGTAAG ACGTGCCTGA TGAAGGCTGT
 901  GCTGAACCTT AAGGACGGGG TCAATGCCTG CATTCTGCCA CTGCTGCAGA
 951  TCGACmGGGA CTCTGGCAAT CCTCAGCCCC TGGTAAATGC CCAGTGCACA
1001  GATGACTATT ACCGAGGCCA CAGCGCTCTG CACATCGCCA TTGAGAAGAG
1051  GAGTCTGCAG TGTGTGAAGC TCCTGGTGGA GAATGGGGCC AATGTGCATG
1101  CCCGGGCCTG CGGCGCAAGC TTATTCCCTT TA
```

Figure 9

```
   1  cTGCACCTGT CGCTGGCTGC CTGCACnAAC CAGnnCCACA TTGTCAACTA
  51  CCTGACGGAG AACCCCCACA AGAAGGCGGA CATGCGGCGC CAGGACTCGC
 101  GAGGCAACAC AGTGCTGCAT GCGCTGGTGG CCATTGCTGA CAACACCCGT
 151  GAGAACACCA AGTTTGTTAC CAAGATGTAC GACCTGCTGC TGCTCAAGTG
 201  TGCnCGCCTC TTCCCCGACA GCAACCTGGA nGCCGTGCTC AACAACGACG
 251  GCCTCTCGCC CCTCATGATG GCTGCCAAGA CGGGCAAGAT TGGGAwCyTT
 301  CAGCACATCA TCCGGCGGGA GGTGACGGAT GAGGACACAC GGCACCTGTC
 351  CCGCAAGTTC AAGGACTGGG CCTATGGGCC AGTGTATTCC TCGCTTTATG
 401  ACCTCTCCTC CCTGGACACG TGTGGGGAAG AGGCCTCCGT GCTGGAGATC
 451  CTGGTGTACA ACAGCAAGAT TGAGAACCGC CACGAGATGC TGGCTGTGGA
 501  GCCCATCAAT GAACTGCTGC GGGACAAGTG GCGCAAgTTC GGGGCCGTCT
 551  CCTTCTACAT CAACGTGGTC TCCTACCTGT GTGCCATGGT CATCTTCACT
 601  CTCACCGCCT ACTACCAGCC GCTGGAGGGC ACACCGCCGT ACCCTTACCG
 651  CACCACGGTG GACTACCTGC GGCTGGCTGG CGAGGTCATT ACGCTCTTCA
 701  CTGGGGTCCT GTTCTTCTTC ACCAACATCA AAGACTTGTT CATGAAGAAA
 751  TGCCCTGGAG TGAATTCTCT CTTCATTGAT GGCTCCTTCC AGCTGCTCTA
 801  CTTCATCTAC TCTGTCCTGG TGATCGTCTC AGCAGCCCTC TACCTGGCAG
 851  GGATCGAGGC CTACCTGGCC GTGATGGTCT TTGCCCTGGT CCTGGGCTGG
 901  ATGAATGCCC TTTACTTCAC CCGTGGGCTG AAGCTgacgg ggacctataG
 951  CATCATGATC CAGAAGATTC TCTTCAAGGA CCTTTTCCGA TTCCTGCTCG
1001  TCTACTTGCT CTTCATGATC GGCTACGCTT CAGCCCTGGT CTCCCTCCTG
1051  AACCCGTGTG CCAACATGAA GGTGTGCAAT GAGGACCAGA CCAACTGCAC
1101  AGTGCCCACT TACCCCTCGT GCCGTGACAG CGAGACCTTC AGCACCTTCC
1151  TCCTGGACCT GTTTAAGCTG ACCATCGGCA TGGGCGACCT GGAGATGCTG
1201  AGCAGCACCA AGTACCCCGT GGTCTTCATC ATCCTGCTGG TGACCTACAT
1251  CATCCTCACC TTTGTGCTGC TCCTCAACAT GCTCATTGCC CTCATGGGCG
1301  AGACAGTGGG CCAGGTCTCC AAGGAGAGCA AGCACATCTG GAAGCTGCAG
1351  TGGGCCACCA CCATCCTGGA CATTGAGCGC TCCTTCCCCG TATTCCTGAG
```

1401 GAAGGCCTTC CGCTCTGGGG AGATGGTCAC CGTGGGCAAG AGCTCGGACG

1451 GCACTCCTGA CCGCAGGTGG TGCTTCAGGG TGGATGAGGT GAACTGGTCT

1501 yACTGGAACC AGAACTTGGG CATCATCAAC GAGGACCCGG GCAAGAATGA

1551 GACCTACCAG TATTATGGCT TCTCGCATAC CGTGGGCCGC CTCCGCAGGG

1601 ATCGCTGGTC CTCGGTGGTA CCCCGCGTGG TGGAACTGAA CAAGAACTCG

1651 AACCCGGACG AGGTGGTGGT GCCTCTGGAC AGCATGGGGA ACCCCCGCTG

1701 CGATGGCCAC CAGCAGGGTT ACCCCCGCAA GTGGAGGACT GATGACGCCC

1751 CGCTCTAGGG ACTGCAGCCC AGCCCCAGCT TCTCTGCCCA CTCATTTCTA

1801 GTCCAGCCGC ATTTCAGCAG TGCCTTCTGG GGTGTCCCCC CACACCCTGC

1851 TTTGGCCCCA GAGGCGAGGG ACCAGTGGAG GTGCCAGGGA GGCCCCAGGA

1901 CCCTGTGGTC CCCTGGCTCT GCCTCCCCAC CCTGGGGTGG GGGCTCCCGG

1951 CCACCTGTCT TGCTCCTATG GAGTCACATA AGCCAACGCC AGAGCCCCTC

2001 CACCTCAGGC CCCAGCCCCT GCCTCTCCAT TATTTATTTG CTCTGCTCTC

2051 AGGAAGCGAC GTGACCCCTG CCCCAGCTGG AACCTGGCAG AGGCCTTAGG

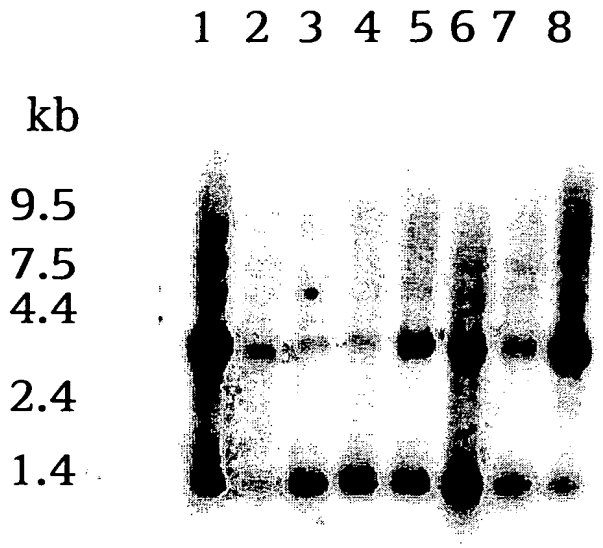
2101 ACCCCGTTCC AAGTGCACTG CCCGGCCAAG CCCCAGCCTC AGCCTGCGCC

2151 TGAGCTGCAT GCGCCACCAT TTTTGGCAGC GTGGCAGCTT TGCAAGGGGC

2201 TGGGGCCCTC GGCGTGGGGC CATGCCTTCT GTGTGTTCTG TAGTGTCTGG

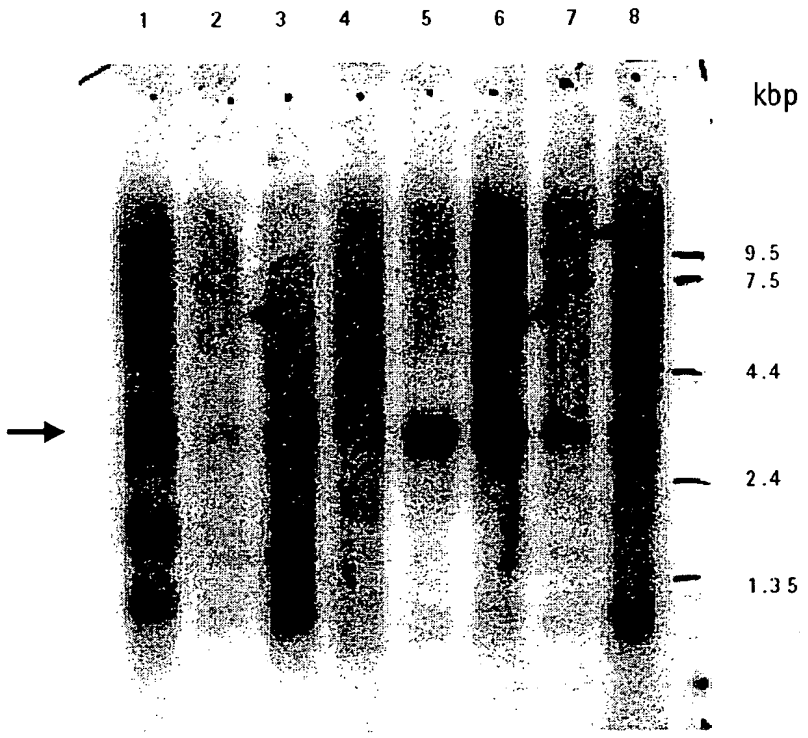
2251 GATTTGCCGG TGCTCAATAA ATGTTTATTC ATTGAAAAAA AAAAAAAAA

Figure 10: Hybridisation of northern blot with hVR2 probe



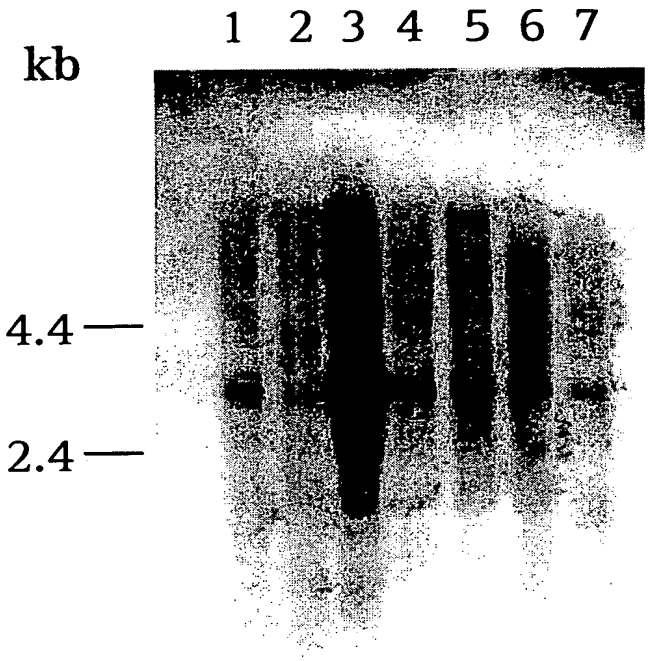
Lane 1: Spleen
Lane 2: Thymus
Lane 3: Prostate
Lane 4: Testis
Lane 5: Ovary
Lane 6: Small intestine
Lane 7: Colon
Lane 8: Peripheral blood leukocyt

Figure 11: Hybridisation of northern blot with hVR2 probe



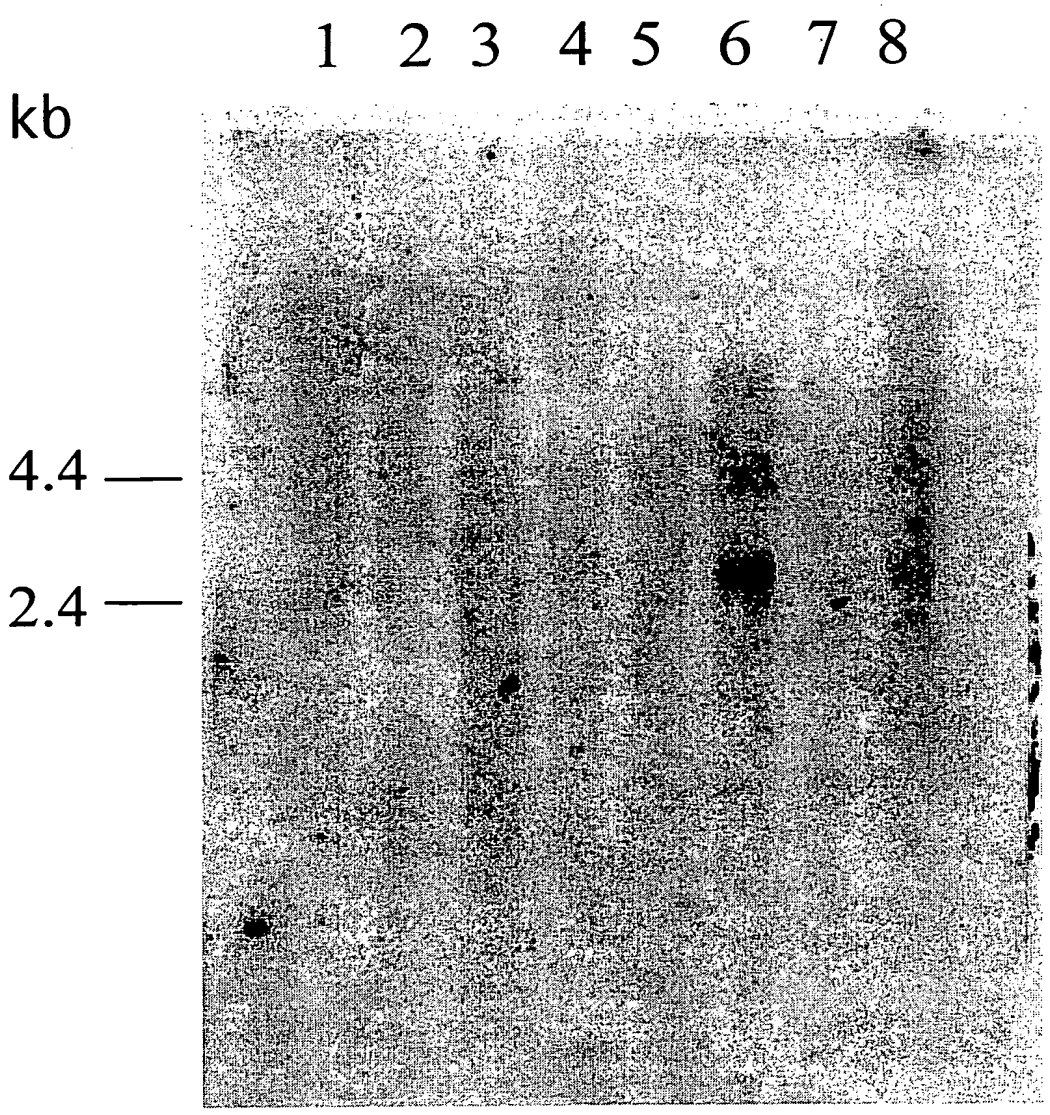
Lane 1: Pancreas
Lane 2: Kidney
Lane 3: Skeletal muscle
Lane 4: Liver
Lane 5: Lung
Lane 6: Placenta
Lane 7: Brain
Lane 8: Heart

Figure 12: Hybridisation of a northern blot with hVR3



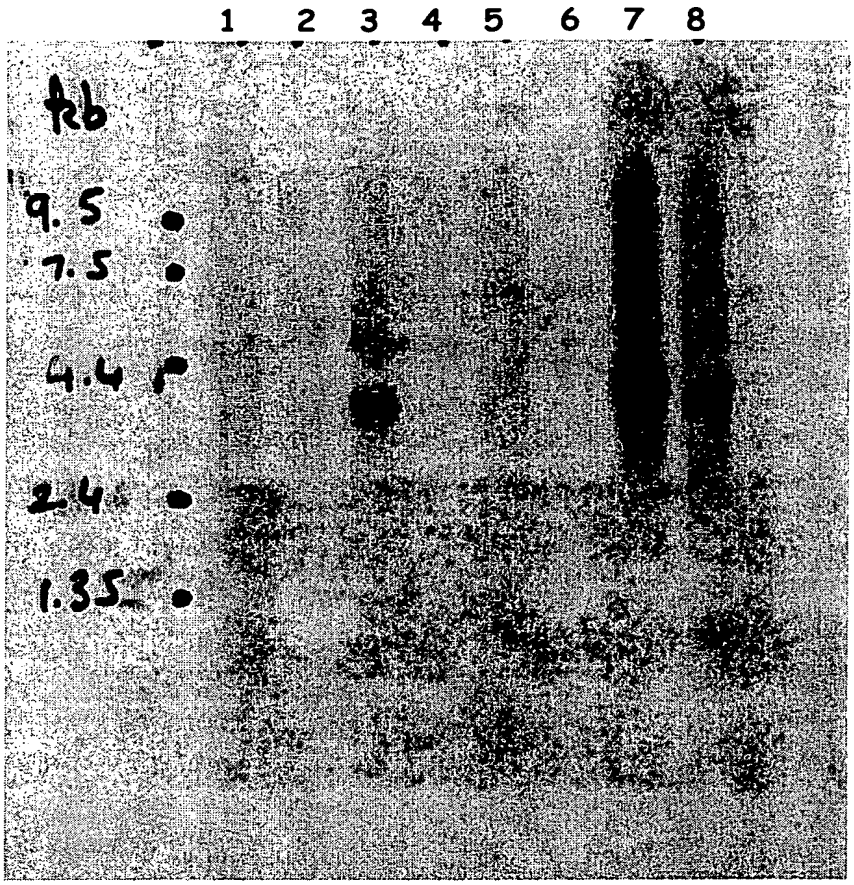
Lane 1: Bone marrow
Lane 2: Adrenal Gland
Lane 3: Trachea
Lane 4: Lymph node
Lane 5: Spinal cord
Lane 6: Thyroid
Lane 7: Stomach

Figure 13: Hybridisation of northern blot with hVR3 probe



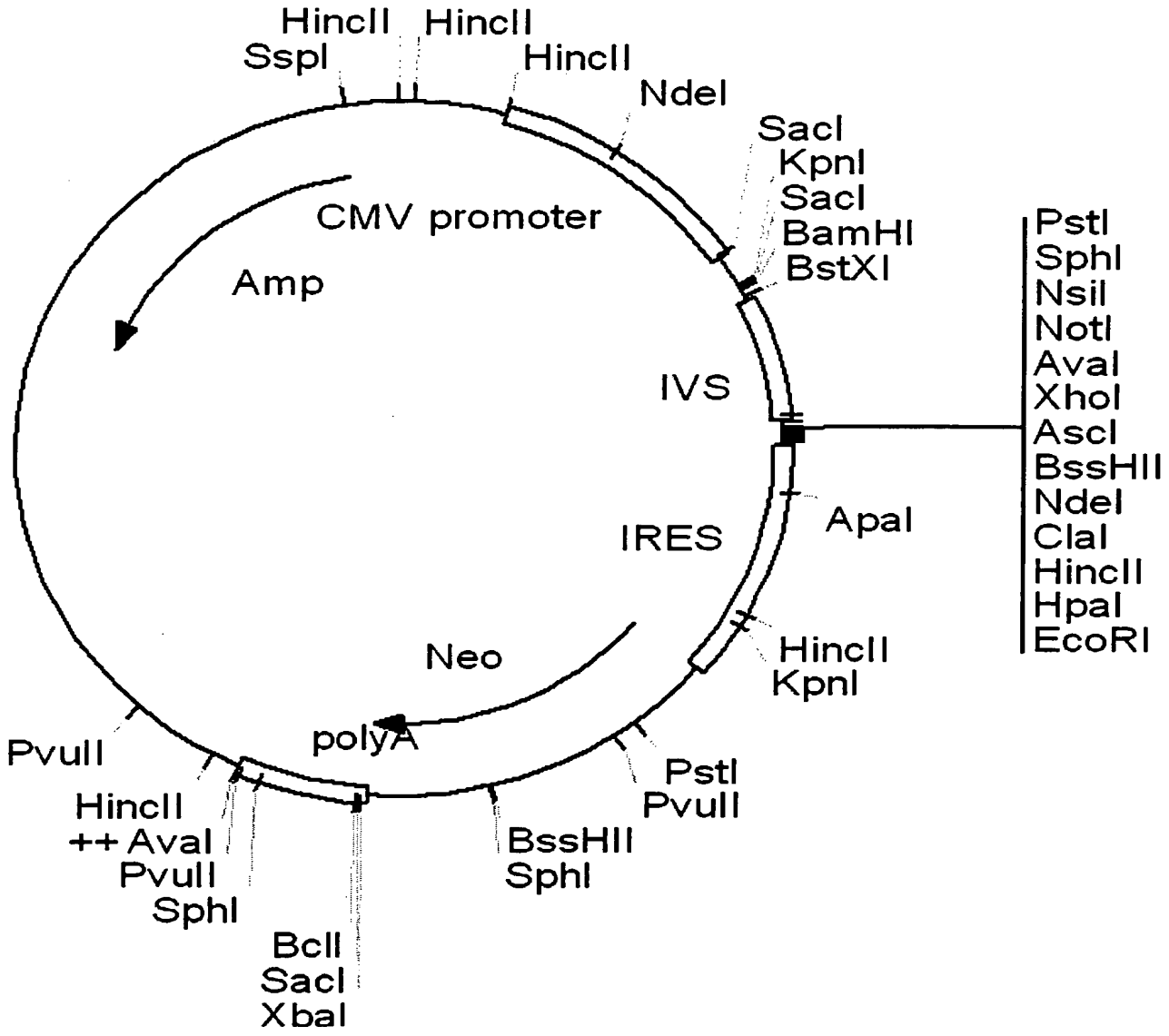
Lane 1: Peripheral Blood Leukocyte
Lane 2: Colon
Lane 3: Small Intestine
Lane 4: Uterus
Lane 5: Testis
Lane 6: Prostate
Lane 7: Thyroid
Lane 8: Spleen

Figure 14: Hybridisation of a multi-tissue northern blot with the hVR3 prob



Lane 1: Heart
Lane 2: Brain
Lane 3: Placenta
Lane 4: Lung
Lane 5: Liver
Lane 6: Skeletal Muscle
Lane 7: Kidney
Lane 8: Pancreas

Figure 15


HincII
HincII
SspI
HincII
NdeI
SacI
KpnI
SacI
BamHI
BstXI
CMV promoter
Amp
IVS
IRES
Neo
polyA
ApaI
HincII
KpnI
PstI
PvuII
BssHII
SphI
BclI
SacI
XbaI
PvuII
HincII
++AvaI
PvuII
SphI
PstI
SphI
NsiI
NotI
AvaI
XhoI
AscI
BssHII
NdeI
ClaI
HincII
HpaI
EcoRI

THIS PAGE BLANK (USPTO)